Package 'sGMRFmix'

October 14, 2022

Type Package

Title Sparse Gaussian Markov Random Field Mixtures for Anomaly Detection

Version 0.3.0

Description An implementation of sparse Gaussian Markov random field mixtures presented by Ide et al. (2016) <doi:10.1109/ICDM.2016.0119>.
It provides a novel anomaly detection method for multivariate noisy sensor data.
It can automatically handle multiple operational modes.
And it can also compute variable-wise anomaly scores.

License MIT + file LICENSE

Encoding UTF-8

LazyData true

RoxygenNote 6.0.1

Imports ggplot2, glasso, mvtnorm, stats, tidyr, utils, zoo

Suggests dplyr, ModelMetrics, testthat, covr, knitr, rmarkdown

VignetteBuilder knitr

NeedsCompilation no

Author Koji Makiyama [cre, aut]

Maintainer Koji Makiyama <hoxo.smile@gmail.com>

Repository CRAN

Date/Publication 2018-04-16 09:27:39 UTC

R topics documented:

compute_anomaly_score *Compute anomaly scores*

Description

Compute anomaly scores

Usage

```
compute_anomaly_score(obj, x, window_size = 1L, ...)
```

Arguments

obj	object
x	data.frame to compute anomaly scores
window_size	integer.
...	additional arguments

Value

matrix of anomaly scores

Examples

```
library(sGMRFmix)

set.seed(314)
train_data <- generate_train_data()
fit <- sGMRFmix(train_data, K = 7, rho = 10)

test_data <- generate_test_data()
compute_anomaly_score(fit, test_data)
```

generate_test_data *Generate test data*

Description

Generate test data

Usage

```
generate_test_data()
```

generate_test_labels *Generate test labels*

Description

Generate test labels

Usage

```
generate_test_labels()
```

generate_train_data *Generate train data*

Description

Generate train data

Usage

```
generate_train_data()
```

plot_multivariate_data *Plot multivariate data*

Description

Plot multivariate data

Usage

```
plot_multivariate_data(df, label = NULL, order_by = index(df),
  guide_title = NULL, fix_scale = FALSE, point_size = 1L)
```

Arguments

df	data.frame of multivariate data
label	data.frame of label for each variables. Or vector of label for each observation.
order_by	vector. An x-axis of plots.
guide_title	character.
fix_scale	logical.
point_size	integer. Point size.

Value

ggplot2 object

Examples

```
library(sGMRFmix)

test_data <- generate_test_data()
test_label <- generate_test_labels()

plot_multivariate_data(test_data)
plot_multivariate_data(test_data, test_label)
```

---

sGMRFmix *Sparse Gaussian Markov Random Field Mixtures*

Description

Sparse Gaussian Markov Random Field Mixtures

Usage

```
sGMRFmix(x, K, rho, kmeans = FALSE, m0 = rep(0, M), lambda0 = 1,
  alpha = NULL, pi_threshold = 1/K/100, max_iter = 500, tol = 0.1,
  verbose = TRUE)
```

Arguments

x	data.frame. A training data.
K	integer. Number of mixture components. Set a large enough number because the algorithm identifies major dependency patterns from the data via the sparse mixture model.
rho	double. Constant that multiplies the penalty term. An optimal value should be determined together with the threshold on the anomaly score, so the performance of anomaly detection is maximized.
kmeans	logical. If TRUE, initialize parameters with k-means method. You should set TRUE for non-time series data. Default FALSE.
m0	a numeric vector. Location parameter of Gauss-Laplace prior. Keep default if no prior information is available. Default 0.
lambda0	double. Coefficient for scale parameter of Gauss-Laplace prior. Keep default if no prior information is available. Default 1.
alpha	double. Concentration parameter of Dirichlet prior. Keep default if no prior information is available. Default 1.
pi_threshold	double. Threshold to decide a number of states. If pi < pi_threshold, the states are rejected in the sense of sparse estimation.

max_iter	integer. Maximum number of iterations.
tol	double. The tolerance to declare convergence.
verbose	logical.

Value

sGMRFmix object

Examples

```
library(sGMRFmix)

set.seed(314)
train_data <- generate_train_data()
fit <- sGMRFmix(train_data, K = 7, rho = 10)
fit
```

Index